

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 25, 2005, 05:40:18 ; Search time 3651 Seconds  
(without alignments)  
14222.654 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272  
1 gaattcggttcacccgcg.....acaccagaagtaagaattc 1272

Sequence: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Scoring table: 45554873 seqs, 2041521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents: NA Main:\*

- 1: /cgn2\_6/prodata/1/pna/PCPUS1\_COMB.seq:\*
- 2: /cgn2\_6/prodata/1/pna/PCPUS2\_COMB.seq:\*
- 3: /cgn2\_6/prodata/1/pna/PCPUS3\_COMB.seq:\*
- 4: /cgn2\_6/prodata/1/pna/US06\_COMB.seq:\*
- 5: /cgn2\_6/prodata/1/pna/US07\_COMB.seq:\*
- 6: /cgn2\_6/prodata/1/pna/US08\_COMB.seq:\*
- 7: /cgn2\_6/prodata/1/pna/US081\_COMB.seq:\*
- 8: /cgn2\_6/prodata/1/pna/US082\_COMB.seq:\*
- 9: /cgn2\_6/prodata/1/pna/US083\_COMB.seq:\*
- 10: /cgn2\_6/prodata/1/pna/US084\_COMB.seq:\*
- 11: /cgn2\_6/prodata/1/pna/US085\_COMB.seq:\*
- 12: /cgn2\_6/prodata/1/pna/US086\_COMB.seq:\*
- 13: /cgn2\_6/prodata/1/pna/US087\_COMB.seq:\*
- 14: /cgn2\_6/prodata/1/pna/US088\_COMB.seq:\*
- 15: /cgn2\_6/prodata/1/pna/US089\_COMB.seq:\*
- 16: /cgn2\_6/prodata/1/pna/US090\_COMB.seq:\*
- 17: /cgn2\_6/prodata/1/pna/US091\_COMB.seq:\*
- 18: /cgn2\_6/prodata/1/pna/US092A\_COMB.seq:\*
- 19: /cgn2\_6/prodata/1/pna/US092B\_COMB.seq:\*
- 20: /cgn2\_6/prodata/1/pna/US093A\_COMB.seq:\*
- 21: /cgn2\_6/prodata/1/pna/US093B\_COMB.seq:\*
- 22: /cgn2\_6/prodata/1/pna/US094\_COMB.seq:\*
- 23: /cgn2\_6/prodata/1/pna/US095A\_COMB.seq:\*
- 24: /cgn2\_6/prodata/1/pna/US095B\_COMB.seq:\*
- 25: /cgn2\_6/prodata/1/pna/US095C\_COMB.seq:\*
- 26: /cgn2\_6/prodata/1/pna/US095D\_COMB.seq:\*
- 27: /cgn2\_6/prodata/1/pna/US096A\_COMB.seq:\*
- 28: /cgn2\_6/prodata/1/pna/US096B\_COMB.seq:\*
- 29: /cgn2\_6/prodata/1/pna/US096C\_COMB.seq:\*
- 30: /cgn2\_6/prodata/1/pna/US096D\_COMB.seq:\*
- 31: /cgn2\_6/prodata/1/pna/US096E\_COMB.seq:\*
- 32: /cgn2\_6/prodata/1/pna/US097A\_COMB.seq:\*
- 33: /cgn2\_6/prodata/1/pna/US097B\_COMB.seq:\*
- 34: /cgn2\_6/prodata/1/pna/US097C\_COMB.seq:\*
- 35: /cgn2\_6/prodata/1/pna/US098A\_COMB.seq:\*
- 36: /cgn2\_6/prodata/1/pna/US098B\_COMB.seq:\*
- 37: /cgn2\_6/prodata/1/pna/US098C\_COMB.seq:\*
- 38: /cgn2\_6/prodata/1/pna/US098D\_COMB.seq:\*
- 39: /cgn2\_6/prodata/1/pna/US099A\_COMB.seq:\*
- 40: /cgn2\_6/prodata/1/pna/US099B\_COMB.seq:\*
- 41: /cgn2\_6/prodata/1/pna/US099C\_COMB.seq:\*
- 42: /cgn2\_6/prodata/1/pna/US099D\_COMB.seq:\*
- 43: /cgn2\_6/prodata/1/pna/US099E\_COMB.seq:\*

- 44: /cgn2\_6/prodata/1/pna/US099F\_COMB.seq:\*
- 45: /cgn2\_6/prodata/1/pna/US099G\_COMB.seq:\*
- 46: /cgn2\_6/prodata/1/pna/US100A\_COMB.seq:\*
- 47: /cgn2\_6/prodata/1/pna/US100B\_COMB.seq:\*
- 48: /cgn2\_6/prodata/1/pna/US101A\_COMB.seq:\*
- 49: /cgn2\_6/prodata/1/pna/US101B\_COMB.seq:\*
- 50: /cgn2\_6/prodata/1/pna/US102A\_COMB.seq:\*
- 51: /cgn2\_6/prodata/1/pna/US102B\_COMB.seq:\*
- 52: /cgn2\_6/prodata/1/pna/US103A\_COMB.seq:\*
- 53: /cgn2\_6/prodata/1/pna/US103B\_COMB.seq:\*
- 54: /cgn2\_6/prodata/1/pna/US104A\_COMB.seq:\*
- 55: /cgn2\_6/prodata/1/pna/US104B\_COMB.seq:\*
- 56: /cgn2\_6/prodata/1/pna/US105A\_COMB.seq:\*
- 57: /cgn2\_6/prodata/1/pna/US105B\_COMB.seq:\*
- 58: /cgn2\_6/prodata/1/pna/US106A\_COMB.seq:\*
- 59: /cgn2\_6/prodata/1/pna/US107A\_COMB.seq:\*
- 60: /cgn2\_6/prodata/1/pna/US107B\_COMB.seq:\*
- 61: /cgn2\_6/prodata/1/pna/US107C\_COMB.seq:\*
- 62: /cgn2\_6/prodata/1/pna/US107D\_COMB.seq:\*
- 63: /cgn2\_6/prodata/1/pna/US108A\_COMB.seq:\*
- 64: /cgn2\_6/prodata/1/pna/US108B\_COMB.seq:\*
- 65: /cgn2\_6/prodata/1/pna/US109A\_COMB.seq:\*
- 66: /cgn2\_6/prodata/1/pna/US109B\_COMB.seq:\*
- 67: /cgn2\_6/prodata/1/pna/US109C\_COMB.seq:\*
- 68: /cgn2\_6/prodata/1/pna/US110\_COMB.seq:\*
- 69: /cgn2\_6/prodata/1/pna/US6000\_COMB.seq:\*
- 70: /cgn2\_6/prodata/1/pna/US6001\_COMB.seq:\*
- 71: /cgn2\_6/prodata/1/pna/US6002\_COMB.seq:\*
- 72: /cgn2\_6/prodata/1/pna/US6003\_COMB.seq:\*
- 73: /cgn2\_6/prodata/1/pna/US6004\_COMB.seq:\*
- 74: /cgn2\_6/prodata/1/pna/US6005\_COMB.seq:\*
- 75: /cgn2\_6/prodata/1/pna/US6006\_COMB.seq:\*
- 76: /cgn2\_6/prodata/1/pna/US6007\_COMB.seq:\*
- 77: /cgn2\_6/prodata/1/pna/US6008\_COMB.seq:\*
- 78: /cgn2\_6/prodata/1/pna/US6009\_COMB.seq:\*
- 79: /cgn2\_6/prodata/1/pna/US6010\_COMB.seq:\*
- 80: /cgn2\_6/prodata/1/pna/US6011\_COMB.seq:\*
- 81: /cgn2\_6/prodata/1/pna/US6012\_COMB.seq:\*
- 82: /cgn2\_6/prodata/1/pna/US6013\_COMB.seq:\*
- 83: /cgn2\_6/prodata/1/pna/US6014\_COMB.seq:\*
- 84: /cgn2\_6/prodata/1/pna/US6015\_COMB.seq:\*
- 85: /cgn2\_6/prodata/1/pna/US6016\_COMB.seq:\*
- 86: /cgn2\_6/prodata/1/pna/US6017\_COMB.seq:\*
- 87: /cgn2\_6/prodata/1/pna/US6018\_COMB.seq:\*
- 88: /cgn2\_6/prodata/1/pna/US6019\_COMB.seq:\*
- 89: /cgn2\_6/prodata/1/pna/US6020\_COMB.seq:\*
- 90: /cgn2\_6/prodata/1/pna/US6021\_COMB.seq:\*
- 91: /cgn2\_6/prodata/1/pna/US6022\_COMB.seq:\*
- 92: /cgn2\_6/prodata/1/pna/US6023A\_COMB.seq:\*
- 93: /cgn2\_6/prodata/1/pna/US6023B\_COMB.seq:\*
- 94: /cgn2\_6/prodata/1/pna/US6024\_COMB.seq:\*
- 95: /cgn2\_6/prodata/1/pna/US6025\_COMB.seq:\*
- 96: /cgn2\_6/prodata/1/pna/US6026\_COMB.seq:\*
- 97: /cgn2\_6/prodata/1/pna/US6027\_COMB.seq:\*
- 98: /cgn2\_6/prodata/1/pna/US6028\_COMB.seq:\*
- 99: /cgn2\_6/prodata/1/pna/US6029\_COMB.seq:\*
- 100: /cgn2\_6/prodata/1/pna/US6030\_COMB.seq:\*
- 101: /cgn2\_6/prodata/1/pna/US6031\_COMB.seq:\*
- 102: /cgn2\_6/prodata/1/pna/US6032\_COMB.seq:\*
- 103: /cgn2\_6/prodata/1/pna/US6033\_COMB.seq:\*
- 104: /cgn2\_6/prodata/1/pna/US6034\_COMB.seq:\*
- 105: /cgn2\_6/prodata/1/pna/US6035\_COMB.seq:\*
- 106: /cgn2\_6/prodata/1/pna/US6036\_COMB.seq:\*
- 107: /cgn2\_6/prodata/1/pna/US6037\_COMB.seq:\*
- 108: /cgn2\_6/prodata/1/pna/US6038\_COMB.seq:\*
- 109: /cgn2\_6/prodata/1/pna/US6039\_COMB.seq:\*
- 110: /cgn2\_6/prodata/1/pna/US6040\_COMB.seq:\*
- 111: /cgn2\_6/prodata/1/pna/US6041\_COMB.seq:\*
- 112: /cgn2\_6/prodata/1/pna/US6042\_COMB.seq:\*
- 113: /cgn2\_6/prodata/1/pna/US6043\_COMB.seq:\*
- 114: /cgn2\_6/prodata/1/pna/US6044\_COMB.seq:\*
- 115: /cgn2\_6/prodata/1/pna/US6045\_COMB.seq:\*
- 116: /cgn2\_6/prodata/1/pna/US6046\_COMB.seq:\*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 25, 2005, 02:16:03 / Search time 1341 Seconds  
(without alignments)  
4061.146 Million cell updates/sec

Title: US-10-077-111-13  
Perfect score: 2047  
Sequence: 1 MVKLITLADHGDPVNCFA.....LTPNRILKMAINRWLETHOK 384

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 26938570 seqs, 7091142127 residues

Total number of hits satisfying chosen parameters: 53877140

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DEV=rlh  
-Q=/cgn2\_1/USPTO\_spool/US10077111/runac\_24102005\_072847\_8258/app\_query.fasta.1.583  
-DB=Pending\_Patents\_NA\_New -QFMT=fastap -SUFFIX=rnpn -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdl  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10077111.OCGN\_1\_141.0runac\_24102005\_072847\_8258 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents\_NA\_New:\*

1: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq2:\*  
3: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq1:\*  
7: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq2:\*  
8: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:\*  
9: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq10:\*  
10: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq11:\*  
11: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq12:\*  
12: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq13:\*  
13: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq14:\*  
14: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq15:\*  
15: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq16:\*  
16: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq17:\*  
17: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq18:\*  
18: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq19:\*  
19: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq20:\*  
20: /cgn2\_6/ptodata/1/pna/US11\_NEW\_COMB.seq21:\*  
21: /cgn2\_6/ptodata/1/pna/US11\_NEW\_COMB.seq22:\*  
22: /cgn2\_6/ptodata/1/pna/US11\_NEW\_COMB.seq23:\*  
23: /cgn2\_6/ptodata/1/pna/US11\_NEW\_COMB.seq24:\*  
24: /cgn2\_6/ptodata/1/pna/US11\_NEW\_COMB.seq25:\*  
25: /cgn2\_6/ptodata/1/pna/US11\_NEW\_COMB.seq26:\*  
26: /cgn2\_6/ptodata/1/pna/US11\_NEW\_COMB.seq27:\*  
27: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

Thu Oct 25 11:58:42 2005

us-10-077-111-12.rnpu

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 25, 2005, 05:49:01 ; Search time 5929 Seconds  
(without alignments)  
3067.275 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272

Sequence: 1 gaattcggtcttaccctgcg.....acacccaagaagtaagaattc 1272

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 27700967 seqs, 7148534824 residues

Total number of hits satisfying chosen parameters: 55401934

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA New:\*

- 1: /cgn2\_6/prodata/2/pna/PCT\_NEW\_COMB.seq2:\*
- 2: /cgn2\_6/prodata/2/pna/PCT\_NEW\_COMB.seq2:\*
- 3: /cgn2\_6/prodata/2/pna/US06\_NEW\_COMB.seq:\*
- 4: /cgn2\_6/prodata/2/pna/US07\_NEW\_COMB.seq:\*
- 5: /cgn2\_6/prodata/2/pna/US08\_NEW\_COMB.seq:\*
- 6: /cgn2\_6/prodata/2/pna/US09\_NEW\_COMB.seq:\*
- 7: /cgn2\_6/prodata/2/pna/US09\_NEW\_COMB.seq1:\*
- 8: /cgn2\_6/prodata/2/pna/US09\_NEW\_COMB.seq2:\*
- 9: /cgn2\_6/prodata/2/pna/US10\_NEW\_COMB.seq:\*
- 10: /cgn2\_6/prodata/2/pna/US10\_NEW\_COMB.seq10:\*
- 11: /cgn2\_6/prodata/2/pna/US10\_NEW\_COMB.seq11:\*
- 12: /cgn2\_6/prodata/2/pna/US10\_NEW\_COMB.seq1:\*
- 13: /cgn2\_6/prodata/2/pna/US10\_NEW\_COMB.seq3:\*
- 14: /cgn2\_6/prodata/2/pna/US10\_NEW\_COMB.seq4:\*
- 15: /cgn2\_6/prodata/2/pna/US10\_NEW\_COMB.seq5:\*
- 16: /cgn2\_6/prodata/2/pna/US10\_NEW\_COMB.seq6:\*
- 17: /cgn2\_6/prodata/2/pna/US10\_NEW\_COMB.seq8:\*
- 18: /cgn2\_6/prodata/2/pna/US10\_NEW\_COMB.seq9:\*
- 19: /cgn2\_6/prodata/2/pna/US11\_NEW\_COMB.seq:\*
- 20: /cgn2\_6/prodata/2/pna/US11\_NEW\_COMB.seq2:\*
- 21: /cgn2\_6/prodata/2/pna/US11\_NEW\_COMB.seq3:\*
- 22: /cgn2\_6/prodata/2/pna/US11\_NEW\_COMB.seq4:\*
- 23: /cgn2\_6/prodata/2/pna/US11\_NEW\_COMB.seq5:\*
- 24: /cgn2\_6/prodata/2/pna/US11\_NEW\_COMB.seq6:\*
- 25: /cgn2\_6/prodata/2/pna/US11\_NEW\_COMB.seq7:\*
- 26: /cgn2\_6/prodata/2/pna/US11\_NEW\_COMB.seq8:\*
- 27: /cgn2\_6/prodata/2/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query			Description
	Score	Match Length	ID	

RP

117: /cgn2\_6/prodata/1/pna/US6047\_COMB.seq:\*  
 118: /cgn2\_6/prodata/1/pna/US6048\_COMB.seq:\*  
 119: /cgn2\_6/prodata/1/pna/US6049\_COMB.seq:\*  
 120: /cgn2\_6/prodata/1/pna/US6050\_COMB.seq:\*  
 121: /cgn2\_6/prodata/1/pna/US6051\_COMB.seq:\*  
 122: /cgn2\_6/prodata/1/pna/US6052\_COMB.seq:\*  
 123: /cgn2\_6/prodata/1/pna/US6053\_COMB.seq:\*  
 124: /cgn2\_6/prodata/1/pna/US6054\_COMB.seq:\*  
 125: /cgn2\_6/prodata/1/pna/US6055\_COMB.seq:\*  
 126: /cgn2\_6/prodata/1/pna/US6056\_COMB.seq:\*  
 127: /cgn2\_6/prodata/1/pna/US6057\_COMB.seq:\*  
 128: /cgn2\_6/prodata/1/pna/US6058\_COMB.seq:\*  
 129: /cgn2\_6/prodata/1/pna/US6059\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	length	ID	Description
------------	-------	-------	-------	--------	----	-------------

102: /cgn2\_6/ptodata/1/pna/US6032\_COMB.seq:\*  
103: /cgn2\_6/ptodata/1/pna/US6033\_COMB.seq:\*  
104: /cgn2\_6/ptodata/1/pna/US6034\_COMB.seq:\*  
105: /cgn2\_6/ptodata/1/pna/US6035\_COMB.seq:\*  
106: /cgn2\_6/ptodata/1/pna/US6036\_COMB.seq:\*  
107: /cgn2\_6/ptodata/1/pna/US6037\_COMB.seq:\*  
108: /cgn2\_6/ptodata/1/pna/US6038\_COMB.seq:\*  
109: /cgn2\_6/ptodata/1/pna/US6039\_COMB.seq:\*  
110: /cgn2\_6/ptodata/1/pna/US6040\_COMB.seq:\*  
111: /cgn2\_6/ptodata/1/pna/US6041\_COMB.seq:\*  
112: /cgn2\_6/ptodata/1/pna/US6042\_COMB.seq:\*  
113: /cgn2\_6/ptodata/1/pna/US6043\_COMB.seq:\*  
114: /cgn2\_6/ptodata/1/pna/US6044\_COMB.seq:\*  
115: /cgn2\_6/ptodata/1/pna/US6045\_COMB.seq:\*  
116: /cgn2\_6/ptodata/1/pna/US6046\_COMB.seq:\*  
117: /cgn2\_6/ptodata/1/pna/US6047\_COMB.seq:\*  
118: /cgn2\_6/ptodata/1/pna/US6048\_COMB.seq:\*  
119: /cgn2\_6/ptodata/1/pna/US6049\_COMB.seq:\*  
120: /cgn2\_6/ptodata/1/pna/US6050\_COMB.seq:\*  
121: /cgn2\_6/ptodata/1/pna/US6051\_COMB.seq:\*  
122: /cgn2\_6/ptodata/1/pna/US6052\_COMB.seq:\*  
123: /cgn2\_6/ptodata/1/pna/US6053\_COMB.seq:\*  
124: /cgn2\_6/ptodata/1/pna/US6054\_COMB.seq:\*  
125: /cgn2\_6/ptodata/1/pna/US6055\_COMB.seq:\*  
126: /cgn2\_6/ptodata/1/pna/US6056\_COMB.seq:\*  
127: /cgn2\_6/ptodata/1/pna/US6057\_COMB.seq:\*  
128: /cgn2\_6/ptodata/1/pna/US6058\_COMB.seq:\*  
129: /cgn2\_6/ptodata/1/pna/US6059\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 25, 2005, 02:08:13 ; Search time 3255.Seconds  
(without alignments)  
4815.990 Million cell updates/sec

Title: us-10-077-111-13

Perfect score: 2047  
Sequence: 1 MVVLHTLADHDVNCACF.....LTPNRLTKAMINRWLETHQK 384

Scoring table:

BLASTSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODES=frame+ p2n.model -DEV=xlh  
-O/cgcn2\_1/USFPO.spool/US10077111/rnrmc\_24102005\_072847\_8228/app.query.fasta.1.583  
-DB=Pending\_Patents\_NA\_Main -QMT=faeCap -SUFFIX=rmnp -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.csl  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10077111 @CGCN 1.1 3250 @rnrmc 24102005 072847 8228 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLDCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA\_Main:

1: /cgcn2\_6/prodata/1/pna/PCTUS1\_COMB.seq:\*  
2: /cgcn2\_6/prodata/1/pna/PCTUS2\_COMB.seq:\*  
3: /cgcn2\_6/prodata/1/pna/PCTUS\_COMB.seq:\*  
4: /cgcn2\_6/prodata/1/pna/US06\_COMB.seq:\*  
5: /cgcn2\_6/prodata/1/pna/US07\_COMB.seq:\*  
6: /cgcn2\_6/prodata/1/pna/US08\_COMB.seq:\*  
7: /cgcn2\_6/prodata/1/pna/US081\_COMB.seq:\*  
8: /cgcn2\_6/prodata/1/pna/US082\_COMB.seq:\*  
9: /cgcn2\_6/prodata/1/pna/US083\_COMB.seq:\*  
10: /cgcn2\_6/prodata/1/pna/US084\_COMB.seq:\*  
11: /cgcn2\_6/prodata/1/pna/US085\_COMB.seq:\*  
12: /cgcn2\_6/prodata/1/pna/US086\_COMB.seq:\*  
13: /cgcn2\_6/prodata/1/pna/US087\_COMB.seq:\*  
14: /cgcn2\_6/prodata/1/pna/US088\_COMB.seq:\*  
15: /cgcn2\_6/prodata/1/pna/US089\_COMB.seq:\*  
16: /cgcn2\_6/prodata/1/pna/US090\_COMB.seq:\*  
17: /cgcn2\_6/prodata/1/pna/US091\_COMB.seq:\*  
18: /cgcn2\_6/prodata/1/pna/US092A\_COMB.seq:\*  
19: /cgcn2\_6/prodata/1/pna/US092B\_COMB.seq:\*  
20: /cgcn2\_6/prodata/1/pna/US093A\_COMB.seq:\*  
21: /cgcn2\_6/prodata/1/pna/US093B\_COMB.seq:\*  
22: /cgcn2\_6/prodata/1/pna/US094\_COMB.seq:\*  
23: /cgcn2\_6/prodata/1/pna/US095A\_COMB.seq:\*  
24: /cgcn2\_6/prodata/1/pna/US095B\_COMB.seq:\*  
25: /cgcn2\_6/prodata/1/pna/US095C\_COMB.seq:\*  
26: /cgcn2\_6/prodata/1/pna/US095D\_COMB.seq:\*  
27: /cgcn2\_6/prodata/1/pna/US096A\_COMB.seq:\*  
28: /cgcn2\_6/prodata/1/pna/US096B\_COMB.seq:\*

29: /cgcn2\_6/prodata/1/pna/US096C\_COMB.seq:\*  
30: /cgcn2\_6/prodata/1/pna/US096D\_COMB.seq:\*  
31: /cgcn2\_6/prodata/1/pna/US096E\_COMB.seq:\*  
32: /cgcn2\_6/prodata/1/pna/US097A\_COMB.seq:\*  
33: /cgcn2\_6/prodata/1/pna/US097B\_COMB.seq:\*  
34: /cgcn2\_6/prodata/1/pna/US097C\_COMB.seq:\*  
35: /cgcn2\_6/prodata/1/pna/US098A\_COMB.seq:\*  
36: /cgcn2\_6/prodata/1/pna/US098B\_COMB.seq:\*  
37: /cgcn2\_6/prodata/1/pna/US098C\_COMB.seq:\*  
38: /cgcn2\_6/prodata/1/pna/US098D\_COMB.seq:\*  
39: /cgcn2\_6/prodata/1/pna/US099A\_COMB.seq:\*  
40: /cgcn2\_6/prodata/1/pna/US099B\_COMB.seq:\*  
41: /cgcn2\_6/prodata/1/pna/US099C\_COMB.seq:\*  
42: /cgcn2\_6/prodata/1/pna/US099D\_COMB.seq:\*  
43: /cgcn2\_6/prodata/1/pna/US099E\_COMB.seq:\*  
44: /cgcn2\_6/prodata/1/pna/US099F\_COMB.seq:\*  
45: /cgcn2\_6/prodata/1/pna/US099G\_COMB.seq:\*  
46: /cgcn2\_6/prodata/1/pna/US100A\_COMB.seq:\*  
47: /cgcn2\_6/prodata/1/pna/US100B\_COMB.seq:\*  
48: /cgcn2\_6/prodata/1/pna/US101A\_COMB.seq:\*  
49: /cgcn2\_6/prodata/1/pna/US101B\_COMB.seq:\*  
50: /cgcn2\_6/prodata/1/pna/US102A\_COMB.seq:\*  
51: /cgcn2\_6/prodata/1/pna/US102B\_COMB.seq:\*  
52: /cgcn2\_6/prodata/1/pna/US103A\_COMB.seq:\*  
53: /cgcn2\_6/prodata/1/pna/US103B\_COMB.seq:\*  
54: /cgcn2\_6/prodata/1/pna/US104A\_COMB.seq:\*  
55: /cgcn2\_6/prodata/1/pna/US104B\_COMB.seq:\*  
56: /cgcn2\_6/prodata/1/pna/US105A\_COMB.seq:\*  
57: /cgcn2\_6/prodata/1/pna/US105B\_COMB.seq:\*  
58: /cgcn2\_6/prodata/1/pna/US106A\_COMB.seq:\*  
59: /cgcn2\_6/prodata/1/pna/US107A\_COMB.seq:\*  
60: /cgcn2\_6/prodata/1/pna/US107C\_COMB.seq:\*  
61: /cgcn2\_6/prodata/1/pna/US107D\_COMB.seq:\*  
62: /cgcn2\_6/prodata/1/pna/US108A\_COMB.seq:\*  
63: /cgcn2\_6/prodata/1/pna/US108B\_COMB.seq:\*  
64: /cgcn2\_6/prodata/1/pna/US109A\_COMB.seq:\*  
65: /cgcn2\_6/prodata/1/pna/US109B\_COMB.seq:\*  
66: /cgcn2\_6/prodata/1/pna/US109C\_COMB.seq:\*  
67: /cgcn2\_6/prodata/1/pna/US110\_COMB.seq:\*  
68: /cgcn2\_6/prodata/1/pna/US110A\_COMB.seq:\*  
69: /cgcn2\_6/prodata/1/pna/US6000\_COMB.seq:\*  
70: /cgcn2\_6/prodata/1/pna/US6001\_COMB.seq:\*  
71: /cgcn2\_6/prodata/1/pna/US6002\_COMB.seq:\*  
72: /cgcn2\_6/prodata/1/pna/US6003\_COMB.seq:\*  
73: /cgcn2\_6/prodata/1/pna/US6004\_COMB.seq:\*  
74: /cgcn2\_6/prodata/1/pna/US6005\_COMB.seq:\*  
75: /cgcn2\_6/prodata/1/pna/US6006\_COMB.seq:\*  
76: /cgcn2\_6/prodata/1/pna/US6007\_COMB.seq:\*  
77: /cgcn2\_6/prodata/1/pna/US6008\_COMB.seq:\*  
78: /cgcn2\_6/prodata/1/pna/US6009\_COMB.seq:\*  
79: /cgcn2\_6/prodata/1/pna/US6010\_COMB.seq:\*  
80: /cgcn2\_6/prodata/1/pna/US6011\_COMB.seq:\*  
81: /cgcn2\_6/prodata/1/pna/US6012\_COMB.seq:\*  
82: /cgcn2\_6/prodata/1/pna/US6013\_COMB.seq:\*  
83: /cgcn2\_6/prodata/1/pna/US6014\_COMB.seq:\*  
84: /cgcn2\_6/prodata/1/pna/US6015\_COMB.seq:\*  
85: /cgcn2\_6/prodata/1/pna/US6016\_COMB.seq:\*  
86: /cgcn2\_6/prodata/1/pna/US6017\_COMB.seq:\*  
87: /cgcn2\_6/prodata/1/pna/US6018\_COMB.seq:\*  
88: /cgcn2\_6/prodata/1/pna/US6019\_COMB.seq:\*  
89: /cgcn2\_6/prodata/1/pna/US6020\_COMB.seq:\*  
90: /cgcn2\_6/prodata/1/pna/US6021\_COMB.seq:\*  
91: /cgcn2\_6/prodata/1/pna/US6022\_COMB.seq:\*  
92: /cgcn2\_6/prodata/1/pna/US6023A\_COMB.seq:\*  
93: /cgcn2\_6/prodata/1/pna/US6023B\_COMB.seq:\*  
94: /cgcn2\_6/prodata/1/pna/US6024\_COMB.seq:\*  
95: /cgcn2\_6/prodata/1/pna/US6025\_COMB.seq:\*  
96: /cgcn2\_6/prodata/1/pna/US6026\_COMB.seq:\*  
97: /cgcn2\_6/prodata/1/pna/US6027\_COMB.seq:\*  
98: /cgcn2\_6/prodata/1/pna/US6028\_COMB.seq:\*  
99: /cgcn2\_6/prodata/1/pna/US6029\_COMB.seq:\*  
100: /cgcn2\_6/prodata/1/pna/US6030\_COMB.seq:\*  
101: /cgcn2\_6/prodata/1/pna/US6031\_COMB.seq:\*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 25, 2005, 04:11:04 ; Search time 3795 Seconds  
(without alignments)  
16241.116 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272

Sequence: 1 gattcggtcttcactcg.....acaccaaagtaagaattc 1272

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hgt:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sta:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1250.2	98.3	1564	6	BC029520 Homo sapi
2	964.2	75.8	1844	9	AR339277 Sequence
3	912.6	71.7	1254	6	CO729621 Sequence
4	893.4	70.2	1919	9	AK129983 Homo sapi
5	890.2	70.0	1811	6	AX748164 Sequence
6	878.8	69.1	1811	9	AK093494 Homo sapi
7	878.8	69.1	1996	6	CQ491312 Sequence
8	878.8	69.1	1996	6	CQ491312 Sequence
9	478.6	37.6	1291	6	AX077675 Sequence
10	474.6	37.3	2130	10	BC061988 Mus muscu
11	436	34.3	1547	10	BC050792 Mus muscu
12	430.2	33.8	3818	5	AJ719461 Gallus ga
13	430	33.8	147021	9	AC006501 Homo sapi
14	430	33.8	177089	9	AC008277 Homo sapi
15	335	26.3	446	6	CQ473162 Sequence
16	331.6	26.1	366	6	CQ096643 Sequence
17	331.6	26.1	366	6	CQ135438 Sequence
18	331.6	26.1	366	6	CQ218732 Sequence
19	331.6	26.1	366	6	CQ294537 Sequence

C 20	331.6	26.1	366	6	CQ311208	CQ311208 Sequence
C 21	329.8	25.9	409	6	CO482331	CO482331 Sequence
C 22	328.8	25.8	441	6	CO503452	CO503452 Sequence
C 23	328.8	25.8	441	6	CO512287	CO512287 Sequence
C 24	297	23.3	297	6	CO109842	CO109842 Sequence
C 25	297	23.3	297	6	CO148540	CO148540 Sequence
C 26	297	23.3	297	6	CO231810	CO231810 Sequence
C 27	297	23.3	297	6	CO307124	CO307124 Sequence
C 28	297	23.3	297	6	CO344013	CO344013 Sequence
C 29	276.8	21.8	245755	2	AC096419	AC096419 Rattus no
C 30	276.8	21.8	315453	2	AC122070	AC122070 Rattus no
C 31	275.4	21.7	196403	2	AC102767	AC102767 Mus muscu
C 32	275.4	21.7	209437	2	AC111119	AC111119 Mus muscu
C 33	275.4	21.7	217352	10	AL929242	AL929242 Mouse DNA
C 34	179.6	14.1	466	6	CO097537	CO097537 Sequence
C 35	179.6	14.1	466	6	CO136375	CO136375 Sequence
C 36	179.6	14.1	466	6	CO219679	CO219679 Sequence
C 37	179.6	14.1	466	6	CO258151	CO258151 Sequence
C 38	179.6	14.1	466	6	CO295485	CO295485 Sequence
C 39	179.6	14.1	466	6	CO332079	CO332079 Sequence
C 40	179.6	14.1	153866	9	AC009307	AC009307 Homo sapi
C 41	176.8	13.9	180	6	CO110706	CO110706 Sequence
C 42	176.8	13.9	180	6	CO149443	CO149443 Sequence
C 43	176.8	13.9	180	6	CO232723	CO232723 Sequence
C 44	176.8	13.9	180	6	CO270754	CO270754 Sequence
C 45	176.8	13.9	180	6	CO308036	CO308036 Sequence

## ALIGNMENTS

RESULT 1	BC029520	1564 bp	mRNA	linear	PRI 20-JUL-2004
LOCUS	BC029520				
DEFINITION	Homo sapiens WD repeat and SAM domain containing 1, mRNA (cDNA clone MGC:33855 IMAGE:5301559), complete cds.				
ACCESSION	BC029520.1	GI:20810486			
VERSION					
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1564)				
AUTHORS	Klausner, R.D., Collins, P.S., Wagner, L., Shemen, C.M., Schler, G.D., Altschul, S.F., Zeeberg, B., Buecaw, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavani, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.U., Mckernan, K.J., Hale, S., Garcia, A.M., Gay, L.J., Huijck, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G., Blakesley, R.W., Touchman, D.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitt, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalys, D.E., Scheraga, A., Schein, J.E., Jones, S.J. and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 1564)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submissions				
JOURNAL	Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA				





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 24, 2005, 21:27:53 ; Search time 3810 Seconds

(without alignments)  
4883.675 Million cell updates/sec

Title: US-10-077-111-13

Perfect score: 2047

Sequence: 1 MVKLHITLADHGDVNCACF.....LTPNRILKMAINRWLETHOK 384

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=x1h  
-O=/cgm2.1/USPTO.spool/US10077111/runat\_24102005\_072845\_8165/app.query.fasta\_1.583  
-DB=GenEmul -OFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10077111 @CGN 1.1 2527 @runat\_24102005\_072845\_8165 -MCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DEPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmul:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2031	99.2	1564	9	BC0295520	Homo sapi
2	1984	96.9	1996	6	CQ491312	Sequence
3	1984	96.9	1996	6	CQ497179	Sequence
4	1978	96.6	1811	6	AX748164	Sequence

Result	Score	Query Match	Length	DB	ID	Description
5	1978	96.6	1811	9	AK093494	Homo sapi
6	1978	96.6	1844	6	AK339277	Sequence
7	1978	96.6	1919	6	AK129983	Homo sapi
8	1823.5	89.1	1254	6	CQ729621	Sequence
9	1617	79.0	2130	10	BC061948	Homo sapi
10	1464	71.5	3818	5	AJ719461	Mus muscu
11	1373	67.1	1547	10	BC050792	Gallus ga
12	1226.5	59.9	1291	6	AX077675	Mus muscu
13	733	35.8	147021	9	AC006501	Homo sapi
14	731	35.7	245795	2	AC096419	Homo sapi
15	731	35.7	315463	2	AC122070	Rattus no
16	728	35.6	177089	9	AC008277	Homo sapi
17	677	33.1	196403	2	AC102767	Mus muscu
18	677	33.1	209437	2	AC111119	Mus muscu
19	677	33.1	217352	10	AL929242	Mouse DNA
20	627.5	30.7	446	6	CQ473162	Sequence
21	626.5	30.6	409	6	CQ482331	Sequence
22	621.5	30.4	441	6	CQ503452	Sequence
23	621.5	30.4	441	6	CQ512287	Sequence
24	597	29.2	366	6	CQ096643	Sequence
25	597	29.2	366	6	CQ135438	Sequence
26	597	29.2	366	6	CQ218732	Sequence
27	597	29.2	366	6	CQ294537	Sequence
28	597	29.2	366	6	CQ331208	Sequence
29	552	27.0	2279	3	AK173854	Ciona int
30	532	26.0	297	6	CQ109842	Sequence
31	532	26.0	297	6	CQ148540	Sequence
32	532	26.0	297	6	CQ231810	Sequence
33	532	26.0	297	6	CQ307124	Sequence
34	532	26.0	297	6	CQ344013	Sequence
35	318	15.5	399	6	AX071746	Homo sapi
36	311	15.2	153866	9	AC009307	Homo sapi
37	308.5	15.1	466	6	CQ097537	Sequence
38	308.5	15.1	466	6	CQ136375	Sequence
39	308.5	15.1	466	6	CQ219679	Sequence
40	308.5	15.1	466	6	CQ258151	Sequence
41	308.5	15.1	466	6	CQ258151	Sequence
42	308.5	15.1	466	6	CQ323079	Sequence
43	300	14.7	180	6	CQ110706	Sequence
44	300	14.7	180	6	CQ149443	Sequence
45	300	14.7	180	6	CQ232723	Sequence

#### ALIGNMENTS

Result	Score	Query Match	Length	DB	ID	Description
1	2031	99.2	1564	9	BC0295520	Homo sapi
2	1984	96.9	1996	6	CQ491312	Sequence
3	1984	96.9	1996	6	CQ497179	Sequence
4	1978	96.6	1811	6	AX748164	Sequence

BC0295520 1564 bp mRNA linear PRI 20-JUL-2004  
Homo sapiens WD repeat and SAM domain containing 1, mRNA (cdna)  
Clone MGC:33855 IMAGE:5301559), complete cds.

BC0295520 1 GI:20810486  
MGC.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 1564)  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Burow, K.H., Scheffer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
Mckernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hult, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smailus, D.B., Scherch, A., Schein, J.E., Jones, S.J. and Narra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 1564)

Strausberg, R.

Direct Submission

Submitted (01-MAY-2002) National Institutes of Health, Mammalian

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [gcaps-remail.nih.gov](mailto:gcaps-remail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

C DNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>

Contact: (Dickson, Mark) [mcdpax1.stanford.edu](mailto:mcdpax1.stanford.edu)

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Series: IRAC Plate: 48 Row: 0 Column: 11

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 22749102.

Location/Qualifiers

1..1564

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:33855 IMAGE:5101559"

/issue\_type="Brain, hypothalamus"

/clone\_11b="N1H MGC\_96"

/lab\_host="DH10B"

/note="Vector: pBluescript"

1..1564

/gene="WDSAM1"

/note="Synonym: FLJ36175"

/db\_xref="locusid:151525"

146..1300

/gene="WDSAM1"

/codon\_start=1

/product="WDSAM1 protein"

/protein\_id="AAH29520.1"

/db\_xref="GI:20810487"

/db\_xref="LocusID:151525"

/translation="MVKLHTLADHDVNCASFSLATGSLDTITLPLYSIDPTE

COFSPDSTCLASGADGTVLVNAGSYKLRGSGYDGLACASPNQSPVRY

GLTITWDDMRCLHSEKADLITCCDSQVSGEQLQFRLASCGDQVIM

VSFTIILARTHOLKQFEDMSFEDVSNLCAQDLKDLVGFKNMIDGKELATK

ESLADLIESIGRSKVRKEEIRTKVSLSSGIPDFICPIRRLMKDPVITASD

YSEKEMENWISKKRISPMNLVLPFAVLPRNTLQKAIIRMLETHOK"

ORIGIN

Alignment Scores:

Pred. No.: 6.72e-177

Score: 2031.00

Percent Similarity: 99.48%

Best Local Similarity: 99.48%

Query Match: 99.22%

Length: 1564

Matches: 382

Conservative: 0

Mismatches: 2

Indels: 0

Gaps: 0

US-10-077-111-13 (1-384) x BC029520 (1-1564)

FEATURES

source

1..1564

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:33855 IMAGE:5101559"

/issue\_type="Brain, hypothalamus"

/clone\_11b="N1H MGC\_96"

/lab\_host="DH10B"

/note="Vector: pBluescript"

1..1564

/gene="WDSAM1"

/note="Synonym: FLJ36175"

/db\_xref="locusid:151525"

146..1300

/gene="WDSAM1"

/codon\_start=1

/product="WDSAM1 protein"

/protein\_id="AAH29520.1"

/db\_xref="GI:20810487"

/db\_xref="LocusID:151525"

/translation="MVKLHTLADHDVNCASFSLATGSLDTITLPLYSIDPTE

COFSPDSTCLASGADGTVLVNAGSYKLRGSGYDGLACASPNQSPVRY

GLTITWDDMRCLHSEKADLITCCDSQVSGEQLQFRLASCGDQVIM

VSFTIILARTHOLKQFEDMSFEDVSNLCAQDLKDLVGFKNMIDGKELATK

ESLADLIESIGRSKVRKEEIRTKVSLSSGIPDFICPIRRLMKDPVITASD

YSEKEMENWISKKRISPMNLVLPFAVLPRNTLQKAIIRMLETHOK"

ORIGIN

Alignment Scores:

Pred. No.: 6.72e-177

Score: 2031.00

Percent Similarity: 99.48%

Best Local Similarity: 99.48%

Query Match: 99.22%

Length: 1564

Matches: 382

Conservative: 0

Mismatches: 2

Indels: 0

Gaps: 0

US-10-077-111-13 (1-384) x BC029520 (1-1564)

FEATURES

source

1..1564

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:33855 IMAGE:5101559"

/issue\_type="Brain, hypothalamus"

/clone\_11b="N1H MGC\_96"

/lab\_host="DH10B"

/note="Vector: pBluescript"

1..1564

/gene="WDSAM1"

/note="Synonym: FLJ36175"

/db\_xref="locusid:151525"

146..1300

/gene="WDSAM1"

/codon\_start=1

/product="WDSAM1 protein"

/protein\_id="AAH29520.1"

/db\_xref="GI:20810487"

/db\_xref="LocusID:151525"

/translation="MVKLHTLADHDVNCASFSLATGSLDTITLPLYSIDPTE

COFSPDSTCLASGADGTVLVNAGSYKLRGSGYDGLACASPNQSPVRY

GLTITWDDMRCLHSEKADLITCCDSQVSGEQLQFRLASCGDQVIM

VSFTIILARTHOLKQFEDMSFEDVSNLCAQDLKDLVGFKNMIDGKELATK

ESLADLIESIGRSKVRKEEIRTKVSLSSGIPDFICPIRRLMKDPVITASD

YSEKEMENWISKKRISPMNLVLPFAVLPRNTLQKAIIRMLETHOK"

ORIGIN

Alignment Scores:

Pred. No.: 6.72e-177

Score: 2031.00

Percent Similarity: 99.48%

Best Local Similarity: 99.48%

Query Match: 99.22%

Length: 1564

Matches: 382

Conservative: 0

Mismatches: 2

Indels: 0

Gaps: 0

US-10-077-111-13 (1-384) x BC029520 (1-1564)

FEATURES

source

1..1564

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:33855 IMAGE:5101559"

/issue\_type="Brain, hypothalamus"

/clone\_11b="N1H MGC\_96"

/lab\_host="DH10B"

/note="Vector: pBluescript"

1..1564

/gene="WDSAM1"

/note="Synonym: FLJ36175"

/db\_xref="locusid:151525"

146..1300

/gene="WDSAM1"

/codon\_start=1

/product="WDSAM1 protein"

/protein\_id="AAH29520.1"

/db\_xref="GI:20810487"

/db\_xref="LocusID:151525"

/translation="MVKLHTLADHDVNCASFSLATGSLDTITLPLYSIDPTE

COFSPDSTCLASGADGTVLVNAGSYKLRGSGYDGLACASPNQSPVRY

GLTITWDDMRCLHSEKADLITCCDSQVSGEQLQFRLASCGDQVIM

VSFTIILARTHOLKQFEDMSFEDVSNLCAQDLKDLVGFKNMIDGKELATK

ESLADLIESIGRSKVRKEEIRTKVSLSSGIPDFICPIRRLMKDPVITASD

YSEKEMENWISKKRISPMNLVLPFAVLPRNTLQKAIIRMLETHOK"

ORIGIN

Alignment Scores:

Pred. No.: 6.72e-177

Score: 2031.00

Percent Similarity: 99.48%

Best Local Similarity: 99.48%

Query Match: 99.22%

Length: 1564

Matches: 382

Conservative: 0

Mismatches: 2

Indels: 0

Gaps: 0

US-10-077-111-13 (1-384) x BC029520 (1-1564)

FEATURES

source

1..1564

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:33855 IMAGE:5101559"

/issue\_type="Brain, hypothalamus"

/clone\_11b="N1H MGC\_96"

/lab\_host="DH10B"

/note="Vector: pBluescript"

1..1564

/gene="WDSAM1"

/note="Synonym: FLJ36175"

/db\_xref="locusid:151525"

146..1300

/gene="WDSAM1"

/codon\_start=1

/product="WDSAM1 protein"

/protein\_id="AAH29520.1"

/db\_xref="GI:20810487"

/db\_xref="LocusID:151525"

/translation="MVKLHTLADHDVNCASFSLATGSLDTITLPLYSIDPTE

COFSPDSTCLASGADGTVLVNAGSYKLRGSGYDGLACASPNQSPVRY

GLTITWDDMRCLHSEKADLITCCDSQVSGEQLQFRLASCGDQVIM

VSFTIILARTHOLKQFEDMSFEDVSNLCAQDLKDLVGFKNMIDGKELATK

ESLADLIESIGRSKVRKEEIRTKVSLSSGIPDFICPIRRLMKDPVITASD

YSEKEMENWISKKRISPMNLVLPFAVLPRNTLQKAIIRMLETHOK"

ORIGIN

Alignment Scores:

Pred. No.: 6.72e-177

Score: 2031.00

Percent Similarity: 99.48%

Best Local Similarity: 99.48%

Query Match: 99.22%

Length: 1564

Matches: 382

Conservative: 0

Mismatches: 2

Indels: 0

Gaps: 0

US-10-077-111-13 (1-384) x BC029520 (1-1564)

FEATURES

source

1..1564

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:33855 IMAGE:5101559"

/issue\_type="Brain, hypothalamus"

/clone\_11b="N1H MGC\_96"

/lab\_host="DH10B"

/note="Vector: pBluescript"

1..1564

/gene="WDSAM1"

/note="Synonym: FLJ36175"

/db\_xref="locusid:151525"

146..1300

/gene="WDSAM1"

/codon\_start=1

/product="WDSAM1 protein"

Db 1226 CCTCAGGGGACTTACACCAATAAGAGCTCTGAAAATGGCCATCAATAGATGGCTGAG 1285  
Qy 381 ThrHsglnlys 384  
Db 1286 ACACACCAAAAG 1297

RESULT 2  
CQ491312 1996 bp DNA linear PAT 30-JAN-2004  
LOCUS Sequence 23179 from Patent WO0160860.  
DEFINITION CQ491312  
ACCESSION CQ491312  
VERSION CQ491312.1 GI:41456931  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1 Schlegel R., Endege W.O. and Monahan J.E.  
Genes differentially expressed in human prostate cancer and their  
use  
Patent: WO 0160860-A 23179 23-AUG-2001;  
Millennium Predictive Medicine, Inc. (US)

FEATURES  
source  
1. 1996  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Alignment Scores:  
Pred. No.: 1,966-172 Length: 1996  
Score: 1984.00 Matches: 383  
Percent Similarity: 80.46% Conservative: 0  
Best Local Similarity: 80.46% Mismatches: 1  
Query Match: 96.92% Indels: 92  
Gaps: 1

US-10-077-111-13 (1-384) x CQ491312 (1-1996)

Qy 1 MetValLysLeuLleHisThrLeuAlaAspHisGlyAspAspValAsnCySerGlyAlaPhe 20  
Db 16 ATGTGGAACATGATTCACACATTAAGCTATCATGTGACGATGTCACATGCTGCTCC 75  
Qy 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrSerLeuArg 40  
Db 76 TCCCTTTCCCTCTGGCTACTGCTCTGACAAAACAATTCGCTGACTGCTTACT 135  
Qy 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaValHisCysCys 60  
Db 136 GACTTTATGAACTGCCAATTCCTCATTTGAAGTTCACTAAGTCTGCTGCTGCTGCTG 195  
Qy 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu 80  
Db 196 TGTTCCTCCCTTCAGCAATATTTTGGCATGCTGCTCAACAGATGCTACCACTGCTCA 255  
Qy 81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100  
Db 256 TGGAAATGTAAGAAAGGACAGATGCTGGAGATGAGAAAGAGCCCTGAGGAGCCCTGG 315  
Qy 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120  
Db 316 AGGCTTCCCACTTTTCCCACTCCACGCTGTTGGCATGAGGAGGAGGAGCTGATGGAAT 375  
Qy 121 ValValLeuTyrPasnAlaGlnSerTyrLysLeuTyrArgCysGlySerValLysAspGly 140  
Db 376 GTGGTTTGTGGAAGGACAGCTCATCAAAATATATAGATGTGATGTTAAAGATGAC 435  
Qy 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160  
Db 436 TCCTTGGGAGGAGTGTGCAATTTCTCCTAATGGAAGCTTTTGTGCACTGGCTCTCACTG 495

Qy 161 GlyAspLeuThrValTyrPasnAspLysMetArgCysIleuHisSerGlyLysAlaHisAsp 180  
Db 496 GGTAATTTAAACAGGTGGATGATTAATAAGTGTGTGCTGATGTAAGAAAAGCAGATGAT 555  
Qy 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200  
Db 556 CTGGAAATTAATCTGCTGCTGATTTTCTTCCACAGCAATTTCTGATGAGAAACAAGCTCT 615  
Qy 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTyrIleValSer 220  
Db 616 CAGTTTTCGACCTGGCATCATGTGTGAGGATTCAGCAAGTCAAAATTTGGATTGTTCT 675  
Qy 221 PheThrHisIleLeu----- 225  
Db 676 TTAAACCATATCTTGAAGTTTGAATTAATAATAAGTACATGAGTGGCACTGTGCT 735  
Qy 225 ----- 225  
Db 736 CCGTCTGCGCTTGCTGCTTTTCCCATGATGGGAGATGCTAGTCTAGGCTCAGTGTGAT 795  
Qy 225 ----- 225  
Db 796 AAGTCTGTCATATATATGATTAATTAATGAGAAATTAATTAATTAATTAATTAATTAAT 855  
Qy 225 ----- 225  
Db 856 ACCAGTATGTCACAACTGCTGCTTTTGCACCTAATACCTTTACTGCTACTGCTTCA 915  
Qy 226 ----- 226  
Db 916 ATGACAAAACAGTGAACATCTGCAATTTGACCTGGAACACTTTGCCAAGCAAGAGGCG 975  
Qy 229 ThrGluHisGlnLeuLysGlnPheThrGluAspTyrSerGluGluValIleSerThrTyr 248  
Db 976 ACAGAAATCAAGCTGGAAGCAATTTACCGAAGATTGTCAGAGGAGAGATGTCTCAACAGG 1035  
Qy 249 LeuCyAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnMetIleAspGly 268  
Db 1036 CTTTGTGCACAAAGATTTAAAGATCTTGTGTGTTTTCAGATGAATTAACATTGATGGA 1095  
Qy 269 LysGluLeuLeuAsnLeuThrLysGlnSerLeuAlaAspAspLysIleGlnSerLeu 288  
Db 1096 AAAGAACTGTTGATCTTCAAAAGAAAGTCTGCTGATGATTTGAAAATTTGAAATCTCTA 1155  
Qy 289 GlyLeuArgSerLysValLeuArgLysIleGlnGluLeuArgThrLysValLysSerLeu 308  
Db 1156 GGACTGGGTAGTAAAGTGTGAGGAAATTAAGAGCTCAGAGCAAGGTTAAATCCCTT 1215  
Qy 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328  
Db 1216 TCTTCAGGAATCTCTGATGATTAATTAATGTCCTCAATTAATGAGAACTTATGAAAGATCCG 1275  
Qy 329 ValIleAlaSerAspGlyTyrSerTyrGluLysGlnLysIleMetGluAsnTyrIleSerLys 348  
Db 1276 GTCAATCGATAGATGGCTATTCATATGAAAGAGAGAGAAATTTGATCAGCAAA 1335  
Qy 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368  
Db 1336 AAGAAGCTACAAATCCCATACAAATCTTCTTCTCAGGGGATCTTACCAAAAT 1395  
Qy 369 ArgThrLeuLysMetAlaIleAsnArgTyrLeuGluThrHisGlnLys 384  
Db 1396 AGGACTCGAAAATGCGCATCAATATAGATGCTGAGACACACCAAAAG 1443

RESULT 3  
CQ497179 1996 bp DNA linear PAT 30-JAN-2004  
LOCUS Sequence 29046 from Patent WO0160860.  
DEFINITION CQ497179  
ACCESSION CQ497179  
VERSION CQ497179.1 GI:41462815  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens